任意のカバー時間を持つ木の構成法 (アルゴリズムと計算機科学の数理的基盤とその応用)

野中, 良哲; 小野, 廣隆; 山下, 雅史

数理解析研究所講究録
第 1691 巻 2010 年 91-95

任意のカバー時間を持つ木の構成法

野中 良哲 * 小野 廣隆 † 山下 雅史 †
* 九州大学大学院システム情報科学府 † 九州大学大学院システム情報科学研究院

概 要

グラフ上のランダムウォークは，粒子がグラフの頂点をある遷移確率に従ってランダムに遷移していくモデルである。n 個の頂点からなる木上のランダムウォークについて、任意のランダムウォークに対するカバー時間が $\Omega(n\log n)$ であることが分かっており，これは星グラフ上の標準ランダムウォークのカバー時間である．また，パスグラフについてはいかなるランダムウォークに対してもカバー時間が $\Omega(n^2)$ であることが知られている．本稿では星グラフとパスグラフとの中間にあるグラフとして箒グラフを定義し，n^2 と $n\log n$ との間の任意のカバー時間のランダムウォークを持つ箒グラフの構造的特徴について考察する．

1 はじめに

グラフ上のランダムウォークとは，グラフ上のある頂点に置かれた粒子がある確率に従ってランダムに遷移していくモデルである．グラフ上のランダムウォークはグラフ $G=(V,E)$ に対して遷移確率行列 P を与えることで定義される。頂点 u から頂点 v へ達するまでの遷移数の期待値を u から v への到達時間という．頂点 v から出発して全ての頂点を訪問するまでの遷移数の期待値を v からのカバー時間という．到達時間およびカバー時間はランダムウォークの速さを測る指標の一つである．

通常単にランダムウォークといった場合は，すべての隣接頂点に等確率で遷移するランダムウォークをさす．このようなランダウォークを標準ランダムウォークと呼ぶ．標準ランダムウォークの到達時間およびカバー時間に関しては，頂点数 n，枝数 m 任意のグラフに対する上界が $O(nm)=O(n^3)$ であることが知られている [1, 2]．

それに対し，池田らは隣接頂点の次数情報を利用したランダムウォークを提案し，到達時間が $O(n^2)$，カバー時間が $O(n^2\log n)$ へそれぞれ改善されることを示している [3]．また，同論文にてパスグラフにおいては，いかなる遷移確率行列に対してもそのカバー時間が $\Omega(n^2)$ であることを示している．この事実は標準ランダムウォークとは異なる遷移確率によるランダムウォーク高速化の可能性と，遷移確率行列を設計することによる高速化の限界が存在することを示している．さて，閉路のない連結グラフである木の中でもっともカバー時間の小さいランダムウォークを持つ木は星グラフであり，その下界が $\Omega(n\log n)$ であることが分かっている [4]．また，一般の木に対し，標準ランダムウォークのカバー時間はすべて $O(n^2)$ かつ $\Omega(n\log n)$ である．本稿では，$O(n^2)$ かつ $\Omega(n\log n)$ である任意の関数 $f(n)$ に対してカバー時間が $\Theta(f(n))$ である木の構造的特徴について考察する．

2 準備

2.1 定義

グラフ上のランダムウォークは，グラフ $G=(V,E)$ に対して遷移確率行列 P を与えることで定義される．トークンはグラフ上のある頂点を出発し，P に従って隣接頂点の一つを選択し，そちらへ遷移するという過程を繰り返す．ランダムウォークの速さを表す指標として到達時間とカバー時間がある．

頂点 s から t への到達時間 $H_G(P;s,t)$ とは，s を出発したトークンが P に従って G 上のランダムウォークを行い．t へ最初に達するまでの遷移数の期待値である．また，その最大値をランダムウォー

クの到達時間 $H_G(P)$ とする．すなわち $H_G(P) = \max_{s,t\in V} H_G(P;s,t)$ である．

頂点 s からの頂点集合 U に対するカバー時間 $C_G(P;s,U)$ とは，s を出発したトークンが P に従って G 上のランダムウォークを行い，U に属する頂点を全て訪問するまでの遷移数の期待値である．また，全頂点に対するカバー時間の最大値をランダムウォークのカバー時間 $C_G(P)$ とする．すなわち $C_G(P) = \max_{s\in V} C_G(P;s,V)$ である．

2.2 枝重みで表現されるランダムウォーク

以降の議論のため，枝重みによるランダムウォークの表現について述べる．各枝 $e=(u,v)$ に重み w_{uv} を与える．枝重みに比例する遷移確率

$$p_{uv} = \frac{w_{uv}}{\sum_{v'\in N(u)} w_{uv'}}. \tag{1}$$

を与えることでランダムウォークを定義することができる．もちろん，全てのランダムウォークが枝重みによって表現できる訳ではない．例えば，ある枝 (u,v) について，$p_{uv} > 0$ かつ $p_{vu} = 0$ であるような遷移確率は表現できない．しかし木上のランダムウォークに限定すれば，次の定理が成り立つ．

定理 2.1 木上の任意のランダムウォークに対して，そのランダムウォークを定義する枝への重み付けが存在する．

証明 木の任意の頂点 r を選ぶ．r に接続する枝の重みを

$$w_{rv} = p_{rv} \qquad v\in N(u) \tag{2}$$

と定義する．この時点で各頂点は以下の 3 つの場合のどれかに必ず当てはまる．

1. 接続するどの枝の重みも決まっていない．
2. ちょうど一つの枝の重みが決定されている．
3. すべての枝の重みが決定済み．

このうち，場合 2 に該当する頂点 u を一つ選ぶ．隣接点の一つ r に接続する枝 w_{ur} の重みが決定されているとし，残りの隣接頂点に接続する枝重みを

$$w_{uv} = \frac{p_{uv}w_{ur}}{p_{ur}} \qquad v\in N(u)\setminus r \tag{3}$$

と決める．今，木を対象としているので，同様の操作を繰り返し適用しても全ての頂点は 3 つの場合のどれかに必ず当てはまり，場合 2 に該当する頂点がなくなったとき，全ての枝重みが決定されている．ここで，u が場合 2 に該当する頂点として選ばれたとき，既に決定されていた唯一の枝重みを w_{ur} とする．この枝重みから得られる u から r への遷移確率 p_{ur} および v への遷移確率 q_{uv} は，

$$q_{uvr} = \frac{w_{ur}}{\sum_{x\in N(u)} p_{ux}w_{ur}/p_{ur}} = p_{rv}, \tag{4}$$

$$q_{uv} = \frac{p_{uv}w_{ur}/p_{ur}}{\sum_{x\in N(u)} p_{ux}w_{ur}/p_{ur}} = p_{uv} \tag{5}$$

となり，もとの遷移確率に等しい． □

2.3 パスグラフ上のランダムウォーク

頂点が一直線に並んだグラフをパスグラフと呼ぶ．

定理 2.2 [3] パス上の任意のランダムウォークのカバー時間は $\Omega(n^2)$ である．

証明 パスグラフ T の頂点を $v_1, v_2, \ldots, v_n$ とし，あるランダムウォークを表現する枝 $(v_i, v_{i+1}), (1\le i < n)$ の重み w_i とする．ランダムウォークの hitting itme は明らかに $\max(H_TP;v_1,v_n, H_T(P;v_n,v_1)$ である．つまり，$H_T(P;v_1,v_n) = H_T(P;v_n,v_1)$ であるようなランダムウォークか到達時間を最小にする．ここで，到達時間と枝重みの関係から，

$$H_T(P;v_1,v_n) + H_T(P;v_n,v_1) = \sum_{i=1}^{l-1}\sum_{j=1}^{l-1} \frac{2w_j}{w_j} \tag{6}$$

である．全枝重みの合計を 1 に固定して考えると，式 6 は全ての枝重みが等しいとき $2(n-1)^2$ で最小になる．カバー時間は当然到達時間以上なので，パスグラフにおけるカバー時間は $\Omega(n^2)$ である．

2.4 木上のランダムウォーク

閉路のない連結グラフを木と呼ぶ．パスグラフは木の一種である．

定理 2.3 木上の任意のランダムウォークのカバー時間は $\Omega(n\log n)$ である．

この定理は以下の 2 つの補題に基づいている．

補題 2.4 木上の任意のランダムウォークに対して，よりカバー時間の小さい星グラフ上のランダムウォークが存在する．

補題 2.5 星グラフ上のランダムウォークで最も高速なものは標準ランダムウォークであり，そのカバー時間は $\Theta(n \log n)$ である．

木上のランダムウォークでは，標準ランダムウォークが

3 箒グラフのカバー時間

3.1 定義

長さ l のパスの一方に k 個の端末点を接続したグラフを箒グラフ $B_{l,k}$ と呼ぶ．箒グラフ $B_{l,k}$ の頂点集合 V および枝集合 E を以下のように定義する．

$$
\begin{aligned}
&T = \{v_1, \ldots, v_l\},\\
&S = \{u_1, \ldots, u_k\},\\
&V = T \cup S,\\
&E = \{(v_i, v_{i+1}) | 1 \leq i < l\} \cup \{(v_l, u_i) | 1 \leq i \leq k\}.
\end{aligned}
$$

図 1 は $l = 4, k = 3$ の箒グラフである．

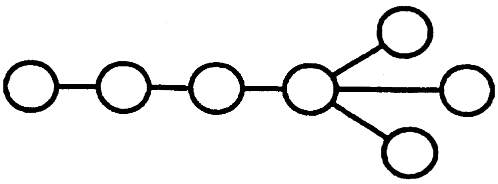

図 1: 箒グラフ $B_{4,3}$

ここで，各枝の重みを，

$$w_{v_i, v_{i+1}} = w_i \qquad (1 \leq i < l), \tag{7}$$

$$w_{v_l, u_i} = w_l / k \qquad (1 \leq i \leq k) \tag{8}$$

と表すことにする．同じ頂点に接続する端末点への遷移確率は，全て等しいときにカバー時間が最小になるので，枝 (v_l, u_i) に与える重みは全て等しいものとする．

3.2 標準ランダムウォーク

定理 3.1 箒グラフ $B_{l,k}$ に対し標準ランダムウォークの遷移確率行列 P_{std} を与える．このとき，$C_{B_{l,k}}(P_{std}) = \Theta(n \log n + nl)$ である．

証明 v_l からの S に対するカバー時間に v_l へ戻るための 1 回の遷移を加えると，

$$
\begin{aligned}
C_{B_{k,l}}(P_{std}; v_l, S) + 1 &= (2k + h_{v_{l-1}, v_l} + 1)h(k),\\
&= 2(n-1)h(k).
\end{aligned} \tag{9}
$$

となる．ただし，$h(k)$ は k 番目の調和関数と呼ばれる関数で，

$$h(k) = \frac{1}{+}\frac{1}{2} + \cdots + \frac{1}{k} \tag{10}$$

である．調和関数 $h(k)$ は漸近的には $\log k$ に等しい．頂点数 $k+1$ の星グラフの中心からのカバー時間に 1 を加えたものが $2kh(k)$ であるので，T 内を移動するために要した遷移数は $(h_{v_{l-1}, v_l} + 1)h(k)$ である．長さ l のパスを端から端まで往復する際の遷移数の期待値は $2(l-1)^2$ であるので，$(h_{v_{l-1}, v_l} + 1)h(k) \geq 2(l-1)^2$ であれば S を全て訪問するする過程で T をすべて訪問することになる．そうでなければ，改めて左端まで行く必要があり，その場合さらに $(n+k-1)(l-1)$ の遷移数を要する．

(i) $(h_{v_{l-1}, v_l} + 1)h(k) \geq 2(l-1)^2$ の場合
$h_{v_{l-1}, v_l} = 2l - 3$ なので，k に関して

$$h(k) \geq l - 1.5 = n - k - 1.5, \tag{11}$$

$$k + h(k) \geq n - 1.5 \tag{12}$$

が得られる，$h(k) < k < n$ なので，これは $k = \Theta(n)$ を意味する．(9) よりカバー時間は $\Theta(n \log n)$ となる．さて，式 (11) から $l = O(\log n)$ であるので，

$$\Theta(n \log k + nl) = \Theta(n \log n) \tag{13}$$

である．

(ii) $(h_{v_{l-1}, v_l} + 1)h(k)2(l-1)^2$ の場合
式 (9) にさらに $2(n+k-1)(l-1)$ が加わるので，カバー時間の見積りとして $\Theta(n \log k + nl)$ が得られる．しかし，条件より $l = \Omega(\log n)$ であるので，結局カバー時間の見積りとしては

$$\Theta(n \log n + nl) = \Theta(nl) \tag{14}$$

である．

パスの途中から出発して v_l に到達するまでの遷移数は高々 $(l-1)^2$ なので，他の出発点からのカバー時間が v_l からのカバー時間より大きいとしてもオーダーとしては変わらない．

定理 3.1 から標準ランダムウォークのカバー時間が任意のオーダーとなる帚グラフを得ることができる．

系 1 $f(n)$ を $f(n) = \Omega(n \log n)$ かつ $f(n) = O(n^2)$ である任意の関数とする．$l = \Theta(f(n)/n)$ である帚グラフについて $C_{B_{l,n-l}}(P_{std}) = \Theta(f(n))$ である．

3.3 最適なランダムウォーク

与えられたグラフに対してカバー時間が最小となるランダムランダムウォークを設計することは一般的には難しいが，帚グラフの場合は比較的簡単に最適なカバー時間を見積もることができる．

定理 3.2 P_{opt} を帚グラフ $B_{l,k}$ に対してカバー時間が最小となる遷移確率行列とする．このとき，$C_{B_{l,k}}(P_{opt}) = \Theta(k \log k + l^2)$ である．

証明 まず，長さ l のパスと k 個の端末点を訪問する必要があることから，

$$C_{B_{l,k}}(P_{opt}) = \Omega(k \log k + l^2) \tag{15}$$

であることは明らかである．次に，実際にカバー時間が $\Theta(k \log k + l^2)$ であるランダムウォークが存在することを示す．

v_l から出発したときのカバー時間を考える．i 個の端末点を訪問済みの状態で v_l に居るとき，新たに $i+1$ 個目の端末点を訪れて v_l に戻るまでの遷移数の期待値を C_i とすると，

$$C_i = \frac{k(1+p+(1-p)H_{v_{l-1},v_l})}{p} \frac{1}{k-i} \tag{16}$$

ただし，$p_{v_l,v_{l-1}} = 1-p$ とし，k 個の端末点への遷移確率は全て等しいとする．S に対するカバー時間 $C_{B_{k,l}}(P; v_l, S)$ に 1 加えた遷移数は

$$\begin{aligned} C_{B_{k,l}}(P; v_l, S) + 1 &= \sum_{i=0}^{k-1} C_i \\ &= \frac{k(1+p+(1-p)H_{v_{l-1},v_l})}{p} h(k) \end{aligned} \tag{17}$$

となる．だたし，$h(k)$ は k 番目の調和関数である．ここで，v_{l-1} から v_l への到達時間を枝重みによって表すと

$$H_{v_{l-1},v_l} = \sum_{j=1}^{l-2} \frac{2w_j}{w_{l-1}} + 1 \tag{18}$$

である．これを代入すると，

$$\begin{aligned} &C_{B_{k,l}}(P_{opt}) \\ &= \frac{2kh(k)}{w_l} = 2kh(k) + \frac{2k(1-w_l)}{w_l} h(k) \end{aligned} \tag{19}$$

となる．このうち，$2kh(k)$ は頂点数 $k+1$ の star グラフのカバー時間+1 そのものであるので，残りがパス部分の移動に費やした遷移数である．ちなみに，パスグラフの往復にかかる遷移数の期待値 H は

$$H = 2 \sum_{i=1}^{l-1} \sum_{i=1}^{l-1} \frac{w_i}{w_j} \tag{20}$$

となる．H は $w_1 = w_2 = \cdots = w_{l-1}$ のとき $2(l-1)^2$ で最小である．

ここで，

$$\frac{2k(1-w_l)}{w_l} h(k) = 2 \sum_{i=1}^{l-1} \sum_{i=1}^{l-1} \frac{w_i}{w_j} \tag{21}$$

となるように w_l を定める．こうすれば，S を全て訪問する過程で T は全て訪問していることになる．このときのカバー時間は，$\Theta(k \log k + l^2)$ である．□

4 まとめ・今後の課題

本稿ではパスグラフと星グラフの中間に位置するグラフとして帚グラフを考え，パラメータを変更することで $\Theta(n^2)$ と $\Theta(n \log n)$ との間にある任意のオーダーのカバー時間が実現できることを示した．今後はより一般的なグラフに対する解析を行い，木におけるランダムウォークのカバー時間がどのような構造的特徴に依存するかを明かにしたいと考えている．

参考文献

[1] R.Aleliunas, R.M Karp, R.J. Lipton, L.Lovaasz, and C.Rackoff, "Random walks, universal traversal sequences, and the complexity of maze problems", Proc. 20th IEEE Ann.Symposium on Foundations of Computer Science, 218-223, 1979.

[2] D.J.Aldous, "On the time taken by random walks on finite groups to visit every state", Z.Wahrsch. verw. Gebiete 62 361-1983.

[3] Satoshi Ikeda, Izumi Kubo, Norihiro Okumoto, Masafumi Yamashita "Impact of Local Topological Information on Random Walks on Finite Graphs", Proceedings of Thirtieth International Colloquium on Automata,Languages and Programming, 1054-1067, 2003.

[4] Yoshiaki Nonaka, Hirotaka Ono, Kunihiko Sadakane, Masafumi Yamashita, "How to design a linear cover time random walk on a finite graph", SAGA 2009.